BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

Microbial Pathogenesis 1993; 15: 327–346

A conserved *Streptococcus pyogenes* extracellular cysteine protease cleaves human fibronectin and degrades vitronectin

Vivek Kapur,[1] Stavros Topouzis,[1] Mark W. Majesky,[1] Ling-Ling Li,[1] Marcie R. Hamrick,[1] Richard J. Hamill,[2] Joseph M. Patti,[3] and James M. Musser[1]*

[1]*Section of Molecular Pathobiology, Department of Pathology, and* [2]*Division of Infectious Diseases, Department of Medicine, and Veterans Administration Medical Center, Baylor College of Medicine, and* [3]*Center for Extracellular Matrix Biology, Institute of Biosciences and Technology, Texas A & M University, Houston, Texas 77030, U.S.A.*

(Received July 19, 1993; accepted in revised form September 6, 1993)

Kapur, V. (Section of Molecular Pathobiology, Department of Pathology, Baylor College of Medicine, Houston, TX 77030, U.S.A.), S. Topouzis, M. W. Majesky, L. L. Li, M. R. Hamrick, R. J. Hamill, J. M. Patti, and J. M. Musser, *Microbial Pathogenesis* 1993, **15**, 327–346.

Streptococcus pyogenes secretes an extracellular cysteine protease that cleaves human interleukin 1β precursor to form biologically active IL-1β, a major cytokine mediating inflammation and shock. To further investigate the potential role of the cysteine protease in host–parasite interactions, the enzyme was purified to apparent homogeneity and tested for ability to degrade several human extracellular matrix proteins. Purified protease cleaved fibronectin, apparently at specific sites, and rapidly degraded vitronectin. In contrast, the protease did not have substantial activity against laminin. The cysteine protease also cleaved fibronectin from human umbilical vein endothelial cells grown *in vitro*. Allelic variation in the cysteine protease structural gene was studied in 67 strains expressing 39 M protein serotypes and five provisional M serologic types, and representing 50 phylogenetically distinct clones identified by multilocus enzyme electrophoresis. The gene is well conserved and allelic variation is due solely to accumulation of point mutations. Based on predicted amino acid sequences, one mature cysteine protease variant would be made by clones expressing serotypes M2, M3, M4, M5, M6, M9, M10, M11, M12, M14, M18, M22, M23, M25, M27, M41, M49, M56, M59, two provisional M types, and two clones non-typeable for M protein. Moreover, 33 of the 39 *speB* alleles identified encode one of three mature protease variants that differ from one another at only one or two amino acids clustered in a ten-amino acid region. All 39 alleles, and virtually all strains, encode a product that reacts with polyclonal antisera specific for purified cysteine protease. No compelling evidence was found for a primitive differentiation of the *speB* gene into two distinct classes, as has been proposed for M protein, opacity factor phenotype, and *vir* regulon architecture. The results demonstrate that the cysteine protease is well conserved in natural populations of *S. pyogenes*, provide additional evidence that this enzyme is involved in host–parasite interactions, and suggest that the protease plays a role in bacterial dissemination, colonization, and invasion, and inhibition of wound healing.

Key words: group A streptococci; extracellular matrix proteins; molecular evolution; multilocus enzyme electrophoresis.

Introduction

Streptococcus pyogenes (group A *Streptococcus*) causes several human diseases, such as pharyngitis, acute rheumatic fever, scarlet fever, poststreptococcal glo-

*Author to whom correspondence should be addressed at: Department of Pathology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, U.S.A.

0882-4010/93/110327+20 $08.00/0

merulonephritis, and toxic-shock-like syndrome.[1] The pathogen makes a variety of putative virulence factors that are believed to aid colonization and contribute to host toxicity, including M protein, C5a peptidase, several immunoglobulin-, fibronectin-, collagen type IV-, and vitronectin-binding proteins, pyrogenic exotoxins A, B, and C, streptokinase, streptolysin O, hyaluronidase, hyaluronic acid, and lipoteichoic acid.[2] Despite considerable research effort, there is no comprehensive understanding of the molecular pathogenesis of group A streptococcal infections. The recent intercontinental increase in streptococcal disease frequency and severity[3,4] has highlighted the need for additional knowledge of bacterium–host interactions at the molecular level.

Several decades ago it was shown that *S. pyogenes* culture supernatants contain a protease that has fibrinolytic activity.[5] The enzyme was subsequently purified,[6] shown to be a cysteine protease,[7] and found to be identical to or an allelic variant of streptococcal pyrogenic exotoxin B (SPE B).[8,9] Considerable circumstantial evidence suggests that the cysteine protease is an important streptococcal virulence factor. For example, Kellner and Robertson[10] demonstrated that intravenous injection of purified cysteine protease causes myocardial necrosis in rabbits. Second, Björck *et al.*[11] found that a cysteine protease inhibitor suppressed *S. pyogenes* growth *in vitro*, and protected mice from lethal bacterial injection. Third, individuals with fatal group A streptococcal infections have lower acute-phase serum antibody levels to SPE B than do patients with less severe infections,[12] a result suggesting a protective role for anti-protease antibody. Fourth, there is evidently significant selective pressure favoring protease production, because with a single exception all *S. pyogenes* strains contain the *speB* gene.[3,13,14] Fifth, recently it was discovered[15] that the cysteine protease cleaves human interleukin 1β precursor to form biologically active IL-1β, a major cytokine mediating inflammation and shock. Sixth, Poon-King *et al.* have reported[16] that a streptococcal extracellular product historically referred to as nephritis-associated protein (NSAP)[17] is identical to all or part of the streptococcal protease precursor. The protein binds to human plasmin, was named nephritis plasmin-binding protein (NPBP), and apparently is involved in the pathogenesis of post-streptococcal glomerulonephritis.[16,17]

Inasmuch as there is accumulating evidence that the cysteine protease is a virulence factor, it is important to further explore its role in host–parasite interactions, and to understand the nature and extent of structural gene allelic variation. We here report that the cysteine protease cleaves human fibronectin (FN) and rapidly degrades vitronectin (VN), results suggesting that the enzyme plays a role in *S. pyogenes* colonization, invasion, and inhibition of wound healing. Comparative sequence analysis identified only limited allelic variation in the SPE B gene in strains recovered from diverse infections, time periods, and geographic localities. Thirty-three of 39 alleles identified in a sample of 67 strains representing 50 clonal genotypes would encode one of three mature protease variants that differ from one another at only one or two amino acids that are clustered in a ten-amino acid region. Group A streptococci harboring all distinct alleles express the SPE B gene.

Results

Cleavage of purified extracellular matrix (ECM) proteins

The results show that the streptococcal cysteine protease rapidly degrades purified VN [Fig. 1(a)]. After 5 min of protease incubation with VN, we were unable to identify degradation products by either Coomassie blue staining or immunoblotting with polyclonal anti-VN antibodies. Similarly, the streptococcal protease cleaved FN immedi-

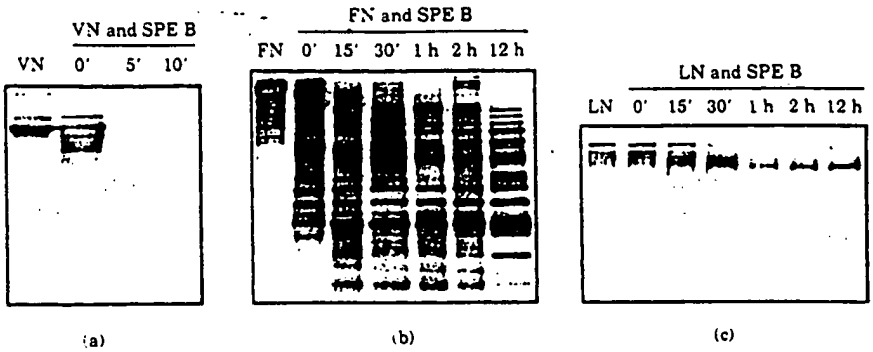

Fig. 1. Cleavage of ECM proteins by streptococcal cysteine protease. Purified streptococcal cysteine protease (0.05 μg) was incubated with either 5 μg of vitronectin (VN) (a), fibronectin (FN) (b), or laminin (LN) (c), at 37°C in PBS (pH 7.4) for the indicated time, and the reaction was stopped by addition of SDS–PAGE sample buffer. The mixture was boiled for 5 min, resolved by SDS–PAGE, transferred to nitrocellulose membranes, and immunoblotted with polyclonal antiserum specific for the ECM protein.

ately, as shown by the rapid appearance of lower molecular weight products [Fig. 1(b)]. However, in contrast to VN degradation, FN cleavage apparently occurred at a limited number of specific sites [Fig. 1(b)]. Incubation of FN with the protease for up to 12 h did not result in formation of additional degradation products.

We did not observe significant cleavage of human laminin (LN) under the experimental conditions assayed [Fig. 1(c)], or when 10 μg of protease and 1 μg of LN substrate are used (data not shown).

Induction of cytopathic effect and fibronectin cleavage in human umbilical vein endothelial cell (HUVEC) cultures

Because patients with invasive *S. pyogenes* episodes frequently have bacterial sepsis with endothelial cell damage, we examined the ability of the streptococcal cysteine protease to cleave FN directly from HUVECs grown in culture. Western immunoblot analysis of cells in the absence of protease, or treated with boiled protease for up to 8 h, showed no detectable FN degradation [Fig. 2(a)]. In contrast, cells incubated with as little as 6 μg/ml of streptococcal protease per well for 2 h retained only a small fraction of intact native FN. Although relatively few time points were examined, the results strongly suggest that the streptococcal protease cleaves FN in a dose-and time-dependent manner in the complex environment of cells growing in tissue culture.

Interestingly, treatment of HUVECs with the streptococcal protease rapidly induced striking cytopathic effects [Fig. 2(b)]. By 3 h after protease addition, zones of clearing occurred in the cell monolayer. This effect was followed by loss of cell adherence to the matrix and ablation of the characteristic cobblestone morphology. We note that FN cleavage was detectable by immunoblot analysis prior to the onset of cytopathic effect.

We did not observe bands that correspond to native human VN in either the solubilized control or treated HUVECs by western immunoblot analysis, presumably due to low level or lack of VN expression by these cells.

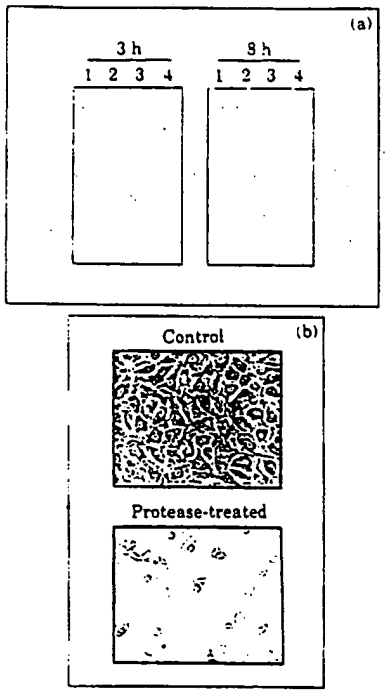

Fig. 2. Effect of streptococcal cysteine protease on HUVECs. (a) Duplicate wells containing HUVECs were either left untreated (lane 1), or treated with PBS (lane 2), boiled streptococcal protease (lane 3), or active streptococcal protease (lane 4) for 3, 5, and 8 h. The contents of the wells were solubilized with SDS–PAGE sample buffer and boiled for 5 min. A 10 μL aliquot was resolved by SDS–PAGE, transferred to a nitrocellulose membrane, and immunoblotted with FN-specific antiserum. The results of the 3 and 8 h incubations are shown. The immunoblot for the 5 h time point is essentially identical to the 8 h time point. (b) Streptococcal protease induces cytopathic effect (CPE) in HUVECs. Duplicate wells containing HUVECs were treated with purified streptococcal cysteine protease and monitored for CPE. There was no apparent difference between the control cells (no protease or addition of boiled protease), and the cells treated with protease for up to 3 h after protease addition. However, considerable CPE was observed after 5 h of incubation with the protease, as demonstrated in the photomicrographs.

Allelic variation in speB: *alleles with solely synonymous substitutions*

Inasmuch as evidence is accumulating that the cysteine protease may be important in the pathogenesis of streptococcal disease, we sought to determine the extent of allelic variation in the protease structural gene. A total of 39 *speB* allelic variants was identified in a sample of 67 strains representing 50 clonal lineages, as indexed by multilocus enzyme electrophoresis (Fig. 3). Compared with the published sequence[9] which we designate *speB1*, the coding region of 11 of the alleles (*speB3*, *speB5*, *speB6*, *speB9*, *speB10*, *speB14*, *speB27*, *speB28*, *speB29*, *speB31*, and *speB35*) was characterized solely by synonymous (silent) nucleotide substitutions (Fig. 4). Among these alleles there were nine polymorphic nucleotide sites in the coding region (positions 205, 483, 765, 819, 913, 915, 993, 999, and 1086 in codons 69, 161, 255, 273, 305, 305, 331, 333, and 362, respectively), eight of which were characterized by T↔C transitions. Strains expressing 14 distinct M types (M1, M2, M3, M4, M5, M6, M12, M18, M22, M23, M25, M41, M49, and M59), two provisional M protein serotypes (PT4854 and PT4931), and three M non-typeable organisms, had a *speB* allele with only silent nucleotide substitutions.

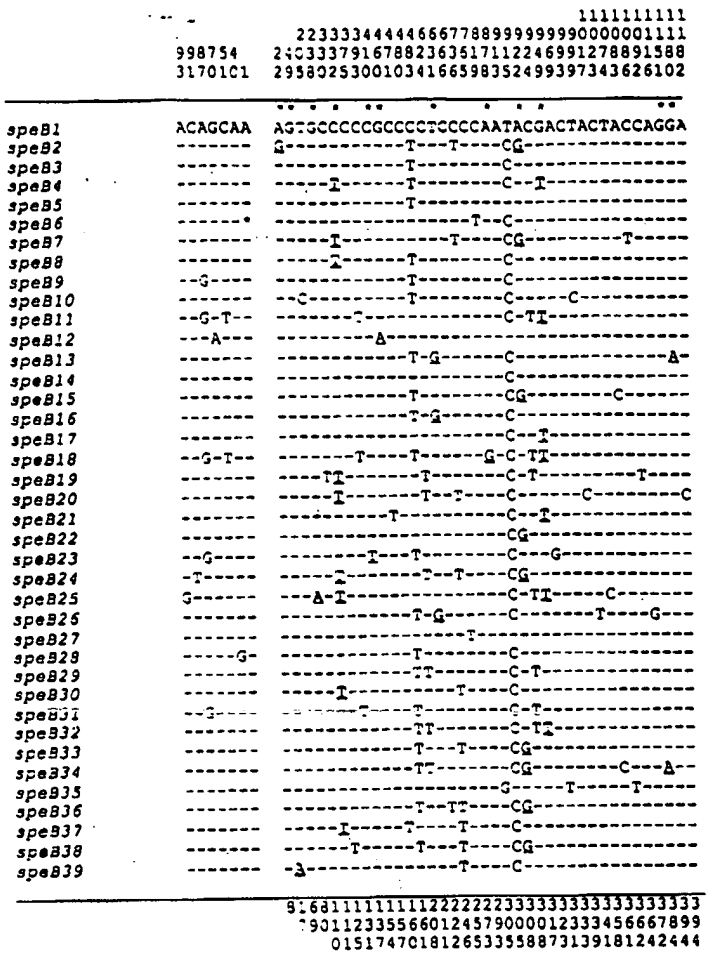

Fig. 3. Alleles of *speB*. The polymorphic sites within the 160 bp upstream noncoding region and 1197 bp coding region of the *speB* gene are shown. The sequence described by Hauser and Schlievert[9] was arbitrarily designated *speB1*, and the numbering of nucleotides and codons is cognate with that sequence. Only those nucleotides in the other alleles that differ from the *speB1* sequence are shown. The position of each polymorphic nucleotide site is shown above the 39 alleles and is numbered in vertical format. Non-synonymous nucleotide changes are underlined and the positions of the coding changes are designated by an asterisk above the coding region polymorphic sites. The seven polymorphic nucleotide sites in the upstream noncoding region are shown in the left of the figure, and the asterisk in *speB6* denotes a deletion of an adenine residue. The codon (numbered in vertical format) containing the polymorphic nucleotide sites is shown below the 39 alleles. The DNA sequence data for *speB2–speB39* are available from EMBL/GenBank/DDBJ under accession numbers L26125–L26162.

Allelic variation in speB: *alleles with non-synonymous substitutions*

Twelve of the 38 polymorphic nucleotide sites result in amino acid replacements in either the zymogen or mature form of the cysteine protease (Fig. 4). Two of the 12 non-synonymous substitutions (in nucleotides 22 and 49) result in amino acid changes in the leader peptide and would not occur in the extracellular zymogen. In addition, three of the 12 non-synonymous substitutions (in nucleotides 238, 332, and 410) are located in the region of the proenzyme that is removed by proteolytic cleavage between Lys-145 and Gln-146,[9] and would not be found in the mature protease. The 39 distinct *speB* alleles would therefore code for a total of 13 different SPE B zymogen molecules

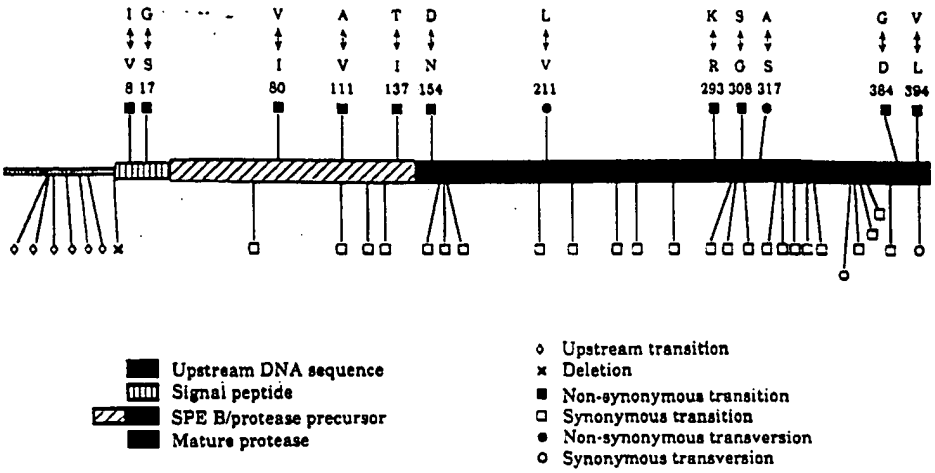

Fig. 4. Schematic representation of polymorphism at the *speB* locus and in the translated protein. The locations of polymorphic sites in 160 bp of upstream DNA and 1197 bp of the *speB* gene identified among 39 alleles in 67 isolates of *S. pyogenes* are shown. There is one deletion of a single adenine residue and 44 nucleotide substitutions, of which 12 result in amino acid replacements. The single letter amino acid abbreviations immediately above the codon numbers refer to residues found in the variant protein[9] arbitrarily designated SPE B1. A, Ala; D, Asp; G, Gly; I, Ile; K, Lys; L, Leu; N, Asn; R, Arg; S, Ser; T, Thr; V, Val.

and eight mature protease variants. A total of 26 distinct alleles code for protein variants with one ($n = 18$ alleles), two ($n =$ seven alleles), or three ($n =$ one allele) amino acid differences compared with *speB1*. Nineteen of the 39 *speB* alleles code for an identical extracellular mature form of the cysteine protease. Similarly, eight alleles (*speB2*, *speB7*, *speB15*, *speB22*, *speB24*, *speB33*, *speB36*, and *speB38*) encode a second identical mature protease, and six alleles (*speB4*, *speB11*, *speB17*, *speB21*, *speB25*, and *speB32*) encode a third mature protease. Therefore, 33 of the 39 *speB* alleles encode one of three mature protease variants that differ from one another at one or two amino acids that are clustered in a ten-amino acid region (amino acid positions 308–317) of the enzyme.

There was a cluster of three amino acid variants in the region between residues 293 and 317 (Fig. 4). Examination of the antigenic index[18] of the protein encoded by *speB* suggests that these amino acids are located in a region of the molecule that is surface exposed and therefore are potentially a target for the host immune response.

There was only a single non-synonymous mutation in nucleotides coding for amino acids 155 through 292, and no non-synonymous substitutions in the gene region specifying amino acids 318 through 383.

Two of the polymorphic amino acids (Ser↔Gly and Ala↔Ser at amino acids 308 and 317, respectively) are located in a region of the molecule that contains a histidine residue believed to be at or near the active site,[19,20] and therefore may alter structure–function characteristics of the mature protease.

Noncoding polymorphic nucleotide sites

Seven nucleotide positions in the upstream non-coding region were polymorphic (Figs 3 and 4). The nucleotides at −93, −91, −87, −70, −51, and −40 had point mutations, and one allele (*speB6*) was characterized by the deletion of one of a series of eight adenine residues located at the −7 to +1 positions. None of the upstream non-coding mutations occurred in the putative −35 and −10 regions or the Shine-Dalgarno sequence.

The nucleotide sequence of the downstream non-coding region was identical in all strains examined and was distinct from the sequence (*speB1*) reported previously.[9] There are five differences that represent apparent insertions and deletions (data not shown).

speB *allele in strain B220 (Elliott strain 5797)*

The *speB* gene (*speB7*) in strain MGAS 1719, which insofar as we can ascertain is the same organism employed by Elliott and co-workers in their studies of streptococcal cysteine protease enzymatic activity and structure,[6] does not encode a protein with the amino acid sequence presented previously.[19,20]

Distribution of speB *alleles and SPE B variants among phylogenetic lineages of* S. pyogenes

In order to gain insight into the pattern of *speB* allele distribution among *S. pyogenes* clonal lineages, we estimated overall genetic relationships among the strains with multilocus enzyme electrophoresis.[3] Comparison of allele profiles over all loci for the 67 strains studied for *speB* variation identified 50 ETs (Table 1), of which 14 (ETs 1, 2, 4, 5, 10, 14, 15, 16, 20, 21, 24, 26, 27, and 32) have been described previously.[3] Estimates of the genetic relationships of the 50 ETs are summarized in the dendrogram in Fig. 5. At a genetic distance of 0.38, there were ten branches, designated A–J. Clusters A, B, D, E, F, and H consisted of 11, 22, 2, 3, 6, and 2 ETs, respectively; and lineages C, G, I, and J were each represented by a single ET.

In the sample of organisms examined, multiple strains of the same combination of multilocus enzyme genotype and M protein serotype had the identical *speB* allele (Table 1). For example, the *speB* sequences of seven ET 1–M1 strains were identical to each other, and, the sequences of four ET 2–M3 organisms were identical.

In contrast, strains expressing the same M serotype, but assigned to different multilocus enzyme genotypes had different *speB* alleles (Fig. 5). Strain MGAS 789, which synthesizes M1 antigen, but is not ET 1, had the *speB5* allele rather than the *speB2* allele found in all other M1 organisms in our sample. Similarly, M12 organisms MGAS 282 (ET 21) and MGAS 590 (ET 24) contained the *speB1* and *speB3* alleles, respectively.

Strains with different multilocus enzyme genotypes can have the identical *speB* allele. For example, six distinct chromosomal lineages (ET 2, ET4, ET 24, ET 32, ET41, and ET57) had the *speB3* allele, strains of ET 20 and ET 21 had the *speB1* allele, strains of ET 14 and ET 37 had the *speB5* allele, strains of ET 26 and 27 had the *speB6* allele, and strains of ET 63 and ET 67 had the *speB24* allele.

speB *alleles and disease type*

Streptococcal clones with the same *speB* allele, and *speB* allele–M protein combination are associated with several different diseases. For example, strains of ET 1–M1–*speB2* were cultured from patients with pharyngitis, scarlet fever, cellulitis, and TSLS; and ET 2–M3–*speB3* organisms were recovered from cases of pharyngitis, scarlet fever, and TSLS. Similarly, strains cultured from individuals with acute rheumatic fever had six distinct *speB* alleles.

SPE B production by S. pyogenes *strains*

The majority of clinical isolates of group A streptococci produce SPE B/cysteine protease[21–23] and patients infected with group A streptococci develop anti-proteinase antibodies.[23] We used immunoblot analysis of culture supernatants to assess production of SPE B/streptococcal protease by strains in our sample of *S. pyogenes*, and one

Table 1 Properties of 68 *S. pyogenes* strains representing 50 ETs[a]

ET	Serotype[b]	MGAS no.[c]	Disease or site[d]	Country and year	*speB* allele	Opacity factor phenotype	M protein class
1	M1	19	Pharyngitis	USA–1980s	*speB2*	–	I
	M1	166	TSLS	USA–1980s	*speB2*		
	M1	285	SID	USA–1980s	*speB2*		
	M1	326	TSLS	USA–1980s	*speB2*		
	M1	480	Invasive	Yugoslavia–1990s	*speB2*		
	M1	579	Cellulitis	Canada–1980s	*speB2*		
	M1	1253	Scarlet fever	UK–1920s	*speB2*		
2	M3	75	Pharyngitis	USA–1980s	*speB3*	–	I
	M3	157	TSLS	USA–1980s	*speB3*		
	M3	315	TSLS	USA–1980s	*speB3*		
	M3	1251	Scarlet fever	USA–1920s	*speB3*		
34	M14	660	Unknown	Egypt–1971	*speB8*	–	I
35	M46	1222	ARF	USA–1953	*speB2*	–	I
36	M76	1832	Unknown	Unknown	*speB33*	+	
21	M12	282	SID	USA–1980s	*speB1*	–	I
37	M1	789	NP	USA–1946	*speB5*	–	I
38	M41	1841	Unknown	Unknown	*speB29*	–	I
39	M33	807	Blood	USA–1969	*speB15*	–	I
40	PT5757	1871	Unknown	Unknown	*speB34*		
41	PT4854	1893	Unknown	Unknown	*speB3*		
42	M8	429	Unknown	Unknown	*speB7*	–	I
43	T8	1719	Unknown	Unknown	*speB7*		
44	M28	587	Scarlet fever	Canada–1980s	*speB18*	–	II
10	T28	289	SID	USA–1980s	*speB11*		
45	M24	684	ARF	USA–1964	*speB12*	–	I
45	M19	1294	ARF	USA–unknown	*speB13*	–	I
46	M29	694	Unknown	Egypt–1971	*speB16*	–	I
46	M44	1226	ARF	USA–1950s	*speB22*	+	II
46	M31	427	Unknown	unknown	*speB16*	–	I
47	M30	366	Unknown	USA–1940s	*speB13*	–	I
26	NT	262	Invasive	USA–1980s	*speB6*		
16	M66	168	Invasive	USA–1980s	*speB25*	+	II
15	M73	302	Invasive	USA–1980s	*speB17*		
14	M4	321	TSLS	USA–1980s	*speB5*	–	II
48	M59	1882	Unknown	Unknown	*speB37*	+	I
49	M5	1289	ARF	USA–1953	*speB14*	–	I
20	M18	156	TSLS	USA–1980s	*speB1*	–	I
	M18	300	Invasive	USA–1980s	*speB1*		
27	M6	303	Invasive	USA–1980s	*speB6*	–	I
50	M43	1842	Unknown	Unknown	*speB36*	–	I
51	M17	1233	ARF	USA–1944	*speB21*	–	I
52	M23	1901	Unknown	Unknown	*speB35*	–	I
53	M49	719	Impetigo	Trinidad–1976	*speB10*	+	II
54	M15	1898	Unknown	Unknown	*speB32*	+	II
55	M25	686	Wound	USA–1969	*speB9*	+	II
24	M12	590	SID	Canada–1980s	*speB3*	–	I
24	M22	162	SID	USA–1980s	*speB28*	+	II
56	M9	800	Impetigo	USA–1964	*speB20*	+	II
57	M56	1864	Unknown	Unknown	*speB3*	–	I
58	M10	1896	Unknown	Unknown	*speB39*		
59	NT	1991	Blood	USA–1993	*speB26*		
	M75	758	ARF	USA–1986	*speB26*		
	M75	1911	Unknown	Unknown	*speB26*		
5	NT	165	TSLS	USA–1980s	*speB4*		
4	M2	327	TSLS	USA–1980s	*speB3*	+	II
60	M9	796	Unknown	USA–1970	*speB19*	+	II
61	M11	650	NP	Trinidad–1972	*speB23*	+	II
	M11	2075	Invasive	Canada–1980s	*speB* null	+	II
62	M62	1883	Unknown	Unknown	*speB17*	+	II

Table 1 *continued.*

ET	Serotype[b]	MGAS no.[c]	Disease or site[d]	Country and year	*speB* allele	Opacity factor phenotype	M protein class
63	M13	659	Unknown	Egypt–1971	*speB24*	+	II
64	PT4931	1870	Unknown	Unknown	*speB31*		
65	TR2612	1872	Unknown	Unknown	*speB38*		
66	M27	1838	Unknown	Unknown	*speB30*	–	I
67	NT	2017	Pharyngitis	USA–1991	*speB24*		
67	NT	2018	Pharyngitis	USA–1992	*speB24*		
68	NT	1990	Pharyngitis	USA–1993	*speB27*		
69	TR2233	1914A	Unknown	Unknown	*speB26*		
32	NT	317	Invasive	USA–1980s	*speB3*		

[a]ET, electrophoretic type.

[b]NT, non-typeable for M protein serotype.

[c]MGAS, Musser group A *Streptococcus* reference number. Strain sources and original designations are as follows: J. C. Huang, Laboratory Centre for Disease Control, Ottawa, Canada, MGAS 579 (11111), 587 (9378), 590 (11078), 2075 (DC11435); J. E. Peters, Wilford Hall Medical Center, San Antonio, TX, MGAS 1991 (BB6672–3), 1990 (BA9812–4), P. M. Schlievert, University of Minnesota, Minneapolis, MN, MGAS 1253 (119/6, also also known as SF130/13), MGAS 1251 (C203S), 166 (Reineke), 285 (195), 325 (89.5.5612), 157 (Zinke), 315 (Soldier 1), 282 (192), 289 (199), 262 (Cal 17), 168 (Reinary), 302 (Lambert), 321 (Weckmuller), 156 (Wilson), 300 (Kluss), 303 (Lundeen), 162 (Cygan), 165 (Wicks), 317 (Timmers); E. L. Kaplan, University of Minnesota, MGAS 480 (90–441); M. A. Kehoe, University of Newcastle-upon-Tyne, Newcastle-upon-Tyne, UK, MGAS 1841 (M41), 1871 (PT5757), 1893 (PT4854), 1882 (M59), 1842 (M43), 1901 (M23), 1898 (M15), 1864 (M56), 1896 (M10), 1911 (M75), 1881 (M62), 1870 (PT4931), 1872 (TR2612), 1838 (M27), 1914A (TR2233); D. LeBlanc, University of Texas Health Science Center at San Antonio, TX, MGAS 1222 (Cole 36XA87), 1226 (Cole 40XF1), 1233 (Cole 45XA9); K. H. Johnston, Louisiana State University Medical Center, New Orleans, LA, MGAS 1719 (B220); D. E. Bessen, Yale University, New Haven, CT, MGAS 1832 (CS110), 1294 (1RP232), 1289 (1RP144); S. K. Hollingshead, Department of Microbiology, University of Alabama School of Medicine, Birmingham, AL, MGAS 660 (D469), 789 (1GL100), 807 (D323), 429 (C256/86/3), 684 (1RP284), 694 (D470), 427 (J137/69/1), 368 (4GL130), 719 (D938), 686 (D316), 800 (A724), 758 (86–809), 796 (D339), 650 (D691), 659 (D474). All other strains are from the collection of J. M. M.

[d]TSLS, toxic-shock-like syndrome; SID, severe invasive disease; ARF, acute rheumatic fever; NP, nasopharynx.

naturally occurring serotype M11 isolate (MGAS 2075) reported to lack *speB*.[13] With the exception of the three strains, all 64 other isolates examined produced cysteine protease, a result consistent with the notion that virtually all *S. pyogenes* strains express the molecule extracellularly (Fig. 6). The three strains had alleles *speB3*, *speB13*, and *speB16*, but other isolates with these same alleles produced the protease (Table 1). Therefore, we conclude that all 39 *speB* alleles can be expressed by group A streptococcal strains under appropriate conditions.

Discussion

Cleavage of fibronectin and degradation of vitronectin

FN is a high-molecular-weight glycoprotein present in plasma and most other body fluids.[24] The molecule is associated with cell surfaces and interstitial connective tissue, and participates in diverse processes such as hemostasis, host defense, and cell adhesion, migration, and differentiation.[25,26] In addition, FN assists in wound healing by mediating migration and attachment of monocytes, fibroblasts, and epithelial cells to the area of injury. FN can also mediate interactions between microorganisms (including *S. pyogenes*[26]), antibody, and leukocytes.[27]

We believe our results may be relevant to several phases of streptococcal pathogenesis. The amount of streptococcal protease produced varies as a function of pH.[8] Under the relatively acidic conditions (~pH 6.0) that occur in tonsillar crypts,[28] protease synthesis increases significantly, and can account for 90% of the protein in

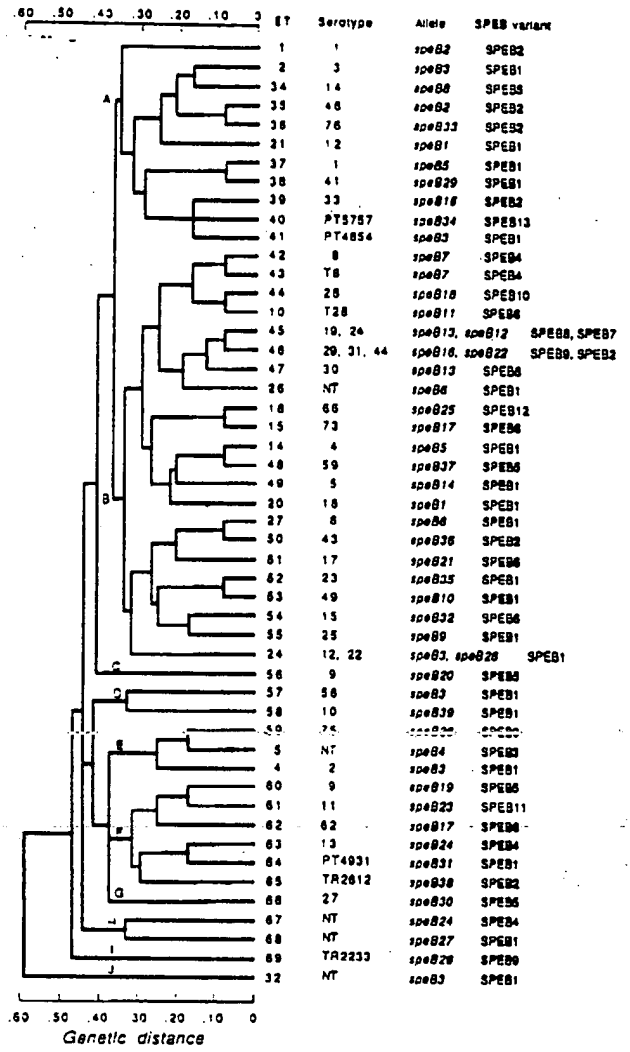

Fig. 5. Dendrogram of 50 *S. pyogenes* ETs. Each isolate was characterized by its combination of alleles at 12 enzyme loci, and distinctive multilocus enzyme genotypes were designated as ETs. The dendrogram was generated from a matrix of genetic distances between pairs of ETs by the average linkage method.[57] Strains are designated by the ET numbers given in Table 1. Unless noted, serotype refers to M protein serotype. PT, provisional M protein serotype; NT, nontypeable for M protein serotype.

streptococcal culture supernatants. It is therefore plausible that the cysteine protease participates in bacterial colonization of the upper respiratory tract through cleavage of ECM proteins, a process that could compromise normal host defenses. Consistent with the idea of protease expression during colonization is the observation that most individuals with streptococcal pharyngitis seroconvert to this molecule.[23]

In principle, the cysteine protease may also be involved in the pathogenesis of deeper, more invasive streptococcal infections, such as cellulitis, osteomyelitis, and toxic-shock-like syndrome (TSLS). Many patients with TSLS have fulminant soft tissue destruction, including fasciitis and myositis, but the pathological basis of the damage is not known. These TSLS characteristics may occur as a consequence of

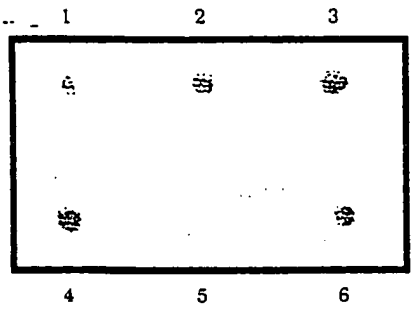

Fig. 6. Western immunoblot assay for expression of SPE B by representative *S. pyogenes* strains. A total of 3 μg of extracellular protein or 0.25 μg of purified protease was spotted onto a nitrocellulose membrane and probed for SPE B with the IgG fraction purified from immune rabbit serum. 1, MGAS 1719; 2, MGAS 166; 3, MGAS 168; 4, MGAS 807; 5, MGAS 2075 (*speB*-negative strain [ref. 13]); and 6, purified protease.

direct tissue degradation by the protease, or the streptococcal enzyme could indirectly trigger deleterious effects on the ECM proteins through activation of human proteases, via conversion of IL-1β precursor to biologically active IL-1β.[15,29] For example, FN cleavage by a variety of proteases has been shown to result in the expression of latent proteolytic activities in FN directed toward ECM proteins.[30]

Gonzalez-Gronow *et al.*[31] have recently shown that streptokinase and human fibronectin share a common epitope that is recognized by both a rabbit and a human anti-streptokinase IgG isolated from the serum of a patient with rheumatoid arthritis. The cross-reactive antibody was identified in the plasma of 40 patients with rheumatoid arthritis, a finding that suggests, but does not prove a causal relationship. Interestingly, the epitope common to these two molecules is apparently not exposed in native fibronectin, rather, it is 'unmasked' as a consequence of digestion by plasmin.[31] Preliminary data suggest that cleavage of fibronectin by the streptococcal cysteine protease exposes the rheumatoid arthritis-associated epitope (Gonzalez-Gronow M., Kapur V. and Musser J. M., unpublished data).

Inasmuch as FN enhances bacterial uptake and killing by polymorphonuclear leukocytes,[25] and mediates binding of group A streptococci to leukocytes,[26] it is possible that cleavage of FN by the cysteine protease enhances bacterial survival *in vivo*. Wound repair in patients infected with group A streptococci may also be detrimentally affected by the cysteine protease. Cleavage of FN by this enzyme *in vivo* might considerably impair the migration and attachment of fibroblasts, epithelial cells, and monocytes that participate in tissue repair and wound resolution.[32]

We note that *Vibrio cholerae* has been shown to express an extracellular protease that cleaves fibronectin and also has hemagglutinin activity.[33] The structural gene (*hap*) for the secreted hemagglutinin/protease has recently been cloned and sequenced, and a null mutant has been obtained.[34] Interestingly, results of experiments using cultured human intestinal cells indicate that the protease is responsible for detachment of the vibrios from the cultured cells by digestion of several putative receptors for *V. cholerae* adhesions.[35] The *V. cholerae* protease data suggest another possible function for the *S. pyogenes* cysteine protease—that of a 'detachase' which permits streptococci to free themselves from host tissues, a process that could facilitate transmission to a new host. In this regard, it is important to note that several groups have described that *S. pyogenes* binds FN,[36-38] and that a streptococcal FN-binding protein enhances bacterial adhesion to respiratory tract epithelial cells.[39,40] It is therefore plausible that expression of the cysteine protease by cells adherent to respiratory epithelium might result in FN cleavage followed by release of bacteria from host tissue. Clearly, additional experi-

ments are required to elucidate the complex interactions between the streptococcal cysteine protease and FN-binding protein, and host FN.

VN is present in plasma, tissues, and extracellular matrices.[41] The molecule has been implicated in several physiological processes, including cell adhesion, and protection of thrombin from inactivation by anti-thrombin III.[41] VN also stabilizes plasminogen activator inhibitor 1 activity, and therefore is important in hemostasis. Our demonstration that the bacterial cysteine protease rapidly degrades VN implies that in the absence of an *in vivo* protease inhibitor, any or all of the normal physiological processes mediated by this protein may be compromised in the course of *S. pyogenes* infection.

Our research adds to data demonstrating that several human pathogenic bacteria express extracellular proteases capable of degrading ECM proteins. Among these organisms are *Pseudomonas aeruginosa*,[42] and *Porphyromonas* (*Bacteriodes*) *gingivalis*,[43,44] two bacterial species that produce host tissue destruction by degradation of collagen and fibronectin, respectively. It is also noteworthy that *Trypanosoma cruzi*, the parasitic flagellate which causes American trypanosomiasis (Chagas' disease), expresses a cell-surface cysteine protease that is a major antigen in humans,[45,46] and is thought to be an important virulence factor.[47] The enzyme (cruzipain) cleaves immunoglobulin G molecules and hydrolyzes the Fc fragment,[46] thereby assisting the organism to evade the immunological consequences of antibody binding. *Entamoeba histolytica*, the cause of amebiasis, also produces an extracellular cysteine protease that is widely believed to be a major virulence factor.[48] Like the streptococcal cysteine protease, the *E. histolytica* enzyme degrades several ECM proteins, including type I collagen, fibronectin, and laminin.[48] The protease also causes cytopathic effect on cell culture monolayers,[49] and is involved with production of tissue necrosis in rat models of acute amebiasis.[50] Interestingly, Reed *et al.*[51] demonstrated that expression of the cysteine protease correlates with the potential of clinical isolates of *E. histolytica* to produce invasive disease. In their study, ten of ten isolates recovered from patients with colitis or amebic liver abscesses expressed the protease, whereas only one of ten organisms cultured from asymptomatic patients made this enzyme. Taken together, the data indicate that there are striking parallels between the functional characteristics of the streptococcal protease and cysteine proteases made by several microbial pathogens.[52]

Inasmuch as individuals with low acute phase antibody levels to SPE B have more severe disease, and are more likely to die,[12] and there is apparently significant pressure to limit *speB* gene diversity in *S. pyogenes*, we favor the hypothesis that the cysteine protease contributes to streptococcal colonization and pathogenesis. Studies are underway to examine the role of the cysteine protease in several phases of streptococcal pathogenesis.

speB *allele in strain B220 (Elliott strain 5797)*

None of the 39 *speB* alleles that we identified, including *speB7* in Elliott strain 5797, would encode a protein with the previously presented amino acid sequence.[19,20] Our data confirm the discrepancies between the protein sequence from strain B220 and a *speB* allele (herein designated *speB1*) in a serotype M12 strain (86–858),[9] and are therefore consistent with the idea advanced earlier[9] that SPE B and streptococcal protease precursor are identical.

speB *alleles and disease type*

In the small sample studied there was no apparent preferential association of *speB* allele and disease type.

The identification of the *speB5* allele in a strain (MGAS 789) recovered in the 1940s

expressing M1 protein, but assigned to ET 37 rather than ET 1 like contemporary M1 strains, suggests that temporal variation in *speB* allele–multilocus enzyme genotype–M protein associations may be a contributing factor in the recent increase in streptococcal disease frequency and severity. Studies are underway to explore this possibility in detail.

speB *variation, M protein class, opacity factor phenotype, and* vir *regulon architecture*

It has been proposed[53,54] that M proteins and their genes can be divided into two major classes based on structural differences. It has also been suggested[55] that the architecture of the *vir* regulon parallels serum opacity factor phenotype and M protein class. Our study found no compelling evidence for an analogous differentiation of *speB* allelic variants. Strains assigned to either of two distinct classes based on reactivity with a panel of monoclonal antibodies to M protein[53] did not have consistent sequence differences, and in several instances the identical *speB* allele was found in strains of two M protein classes. For example, we found that the *speB3* allele occurred in strains of both class I (M3 and M12) and class II (M2), and similarly, the *speB5* allele was identified in strains expressing M1 and M4 assigned to class I and class II, respectively (Table 1). Similarly, there was no simple congruent relationship between *speB* allele and *vir* regulon architecture or opacity factor phenotype. M2, M3, and M12 strains all had the *speB3* allele, but, M3 and M12 are opacity factor-negative and M2 is opacity factor-positive (Table 1). Therefore, if differentiation of *S. pyogenes* strains into two distinct lineages now marked by M protein class, opacity factor phenotype, and *vir* regulon architecture was a primitive event in the evolution of the species,[55] our data can be interpreted to mean that the divergence occurred prior to acquisition of the *speB* gene. We note that the lack of a significant correlation between M serotype class and *speB* phylogeny could also be caused by relatively frequent lateral transfer events involving part or all of the *emm* and *speB* genes.

Evidence for horizontal transfer of speB

Although most of the *speB* allele–clone associations we identified can be explained by chromosomal divergence occurring concurrently with or after the differentiation of the *speB* alleles, there is evidence suggesting that the gene is occasionally laterally transferred in natural populations. The clearest example is the association of the *speB3* allele with strains of ET 2 and ET 32, which diverge from one another at a genetic distance ($\sim$0.55) that in many bacteria approximates or exceeds species-level distinction.[56,57] Other examples of possible horizontal transfer events include the allele *speB5* which is found in lineages A and B, and the *speB24* allele which is found in lineages F and H (Fig. 5).

The identification of molecular population genetic evidence consistent with horizontal spread of the cysteine protease gene between clones of *S. pyogenes* is noteworthy because there has been debate regarding the possibility that *speB* is bacteriophage-associated. Colon-Whitt *et al.*[58] and Johnson *et al.*[59] failed to identify evidence of transfer of SPE B expression by lysogenic conversion, but a later study[60] of a relatively large number of strains reported conversion to SPE B production by phages obtained from organisms of several M protein serotypes. We favor the idea that occasional episodes of *speB* horizontal transfer and recombination (as initially suggested by the phage conversion studies) have played a role in the evolution of *S. pyogenes* clone–*speB* allele associations.

Population genetic inferences and codon utilization bias

There has long been evidence that alternative synonymous codons are not functionally equivalent,[61,62] and in recent years, DNA sequencing studies have demonstrated that

for many bacterial genes the alternative synonymous codons for each amino acid are not used randomly. It has also been shown that within species there is heterogeneity among genes in the use of alternative synonymous codons, and that there is frequently a positive correlation between degree of codon bias and level of gene expression.[63–66] In addition, discrete patterns of codon usage in coliphage T4 genes are related to the time of gene expression.[61] For example, the early phage genes have a codon usage that approximates that of the host *Escherichia coli*, whereas the late genes exhibit a shift in preference toward those codons recognized by the phage-encoded tRNAs. Inasmuch as preliminary analysis of codon utilization for *speB* and other streptococcal pyrogenic exotoxin genes demonstrated a striking level of codon bias for *speB*, and the level of *speB* expression is very high in many strains,[8] we examined the patterns of codon usage in *S. pyogenes* virulence genes deposited in GenBank. The analysis found a high level of codon bias for *speB* compared to *speA*, *speC*, and virtually all other *S. pyogenes* genes analyzed. For example, there is preferential use of AAC and AAT for Asn (G-sq. = 19.52), TAC over TAT for Tyr (G-sq. = 4.31), and AAA over AAG for Lys (G-sq. = 11.13).

Our results are consistent with the idea that nucleotide variation in this gene is constrained among isolates of the species in part because of strong codon utilization bias. This hypothesis is particularly attractive because it provides a plausible explanation for the relative paucity of synonymous substitutions in *speB*. One practical implication of the observation of codon utilization bias is that it suggests that as more *speB* alleles are characterized, they will be largely nominal variants of those we describe here.

Potential immunoprophylaxis ramifications

An efficacious vaccine is not yet available for *S. pyogenes*, in part because of the considerable structural and antigenic diversity present in M protein,[67] an antiphagocytic molecule that has been the focus of virtually all immunoprophylaxis research efforts. The observation that the cysteine protease is well conserved in naturally occurring clones responsible for most disease episodes and recovered from intercontinental sources decades apart, together with several lines of evidence that the protease is involved in the pathogenesis of *S. pyogenes* infections, may have significant implications for vaccine research. Although we have thus far examined sequence diversity in only about one-half of the more than 80 M protein serotypes, our sample was chosen to include strains representing the breadth of multilocus enzyme genotypic diversity present in the species *S. pyogenes* (unpublished data), and most M protein serotypes commonly recovered from patients with pharyngitis or several more serious diseases. This strategy was designed to enhance the likelihood of identifying highly variant *speB* alleles occurring in natural populations. Our work clearly shows that *speB* diversity is considerably restricted compared to the M protein structural gene (*emm*).[67]

The level of *speB* allelic diversity also appears to be less than that described for the *scpA* gene,[68] coding for streptococcal C5a peptidase, a molecule that may be of immunoprophylaxis research interest.[69] Cleary *et al.*[70] recently reported that the *scpA49* allele is 1.5 kb larger than the *scpA12* allele and suggested that *scpA49* encodes a protein with a peptide insert with antigenic epitopes lacking in SCPA12. We have not detected size variation in PCR products of more than 100 *S. pyogenes* strains not yet sequenced for *speB*, and our data show that two clones expressing M12 and one synthesizing M49 differ in *speB* sequence by only a few synonymous nucleotide changes. In regard to employing C5a peptidase in immunoprophylaxis studies, we note that a peptide motif (-N-X-X-X-X-Q-F-Q-) associated with autoimmune disease of the ovary was recently described;[71] this motif is also present in the SCP molecule

(amino acids 828 to 835 in ref. 68). Inasmuch as several streptococcal diseases are thought to have an autoimmune component, we believe this finding is of potential concern in the context of utilization of inactivated C5a peptidase as a human vaccine.

Conclusions

Our data demonstrate that an extracellular cysteine protease made by the pathogenic bacterium *S. pyogenes* cleaves human FN and degrades VN. The gene is well conserved in streptococcal isolates recovered from diverse clinical syndromes, geographic localities, and time periods. The cysteine protease has also been shown recently to cleave human IL-1β precursor to form biologically active IL-1β, a major cytokine mediating inflammation and shock.[15] We believe these data, taken together with the results of previous studies[5,11,12,16] clearly suggest a role for the cysteine protease in one or more phases of streptococcal pathogenesis.

Materials and methods

Bacterial isolates. A sample of 68 strains of *S. pyogenes* (Table 1) was studied. MGAS 1719 is identical to strain B220 and was kindly provided by Dr K. H. Johnston, Louisiana State University Medical School, New Orleans, LA. The strain was originally obtained from the Lancefield collection at Rockefeller University; B220 is the designation assigned by Dr R. Lancefield to strain 5797 studied by S. D. Elliott (D. Bessen, personal communication, and see ref. 72). The strain expresses type 8 T antigen but is serologically nontypeable for M protein.[72] Insofar as we can ascertain, strains B220 and 5797 share lineal descent from a recent common precursor. Strain MGAS 2075, a naturally occurring *speB*-negative organism,[13] was kindly supplied by Dr J. C. Huang, Laboratory Centre for Disease Control, Ottawa, Canada. The proximal sources of the strains are listed in a footnote to Table 1.

Purification of streptococcal cysteine protease. The cysteine protease was purified from *S. pyogenes* strain MGAS 1719 by a procedure described elsewhere.[15] Briefly, bacteria were grown in a chemically defined medium (JRH Biosciences, Lenexa, KS) maintained at pH 5.5–6.0, and the protease was purified from the concentrated culture supernatant by column chromatography with matrex gel red A (Amicon, MA).

SDS–polyacrylamide gel electrophoresis and Coomassie blue staining of the resulting proteolytically active material showed a single major band with apparent M, ~30 kDa.[16] Sequence analysis of the first ten aminoterminal residues (QPVVKSLLDS) with an Applied Biosystems, Inc. model 477A protein sequencer (Protein Sequencing Facility, Baylor College of Medicine) confirmed the identity of the purified material as streptococcal cysteine protease.[9,15]

Detection of extracellular matrix (ECM) protein cleavage by western immunoblot analysis. To test for cleavage of purified human ECM proteins, 5 μg of vitronectin (VN), fibronectin (FN), or laminin (LN) (Calbiochem) suspended in phosphate buffered saline, pH 7.4 (PBS), were incubated with 0.05 μg–0.5 μg streptococcal protease at 37°C for varying times. The reaction was stopped by the addition of an equal volume of 2X SDS–PAGE sample buffer (125 mM Tris-HCl, pH 6.8, 20% glycerol, 4.1% SDS, 2% 2-mercaptoethanol, and 0.001% bromophenol blue) followed immediately by boiling for 5 min. The mixtures were resolved by SDS–PAGE and transferred to a nitrocellulose membrane (Hybond ECL, Amersham, IL). The unreacted membrane sites were blocked by incubation with 0.5% blocking agent (Amersham, IL) for 1 h. The membrane was rinsed with PBS-Tween 20 (0.05%), and incubated for 30 min with a 1:5000 dilution of rabbit polyclonal antiserum raised against the test ECM protein (Gibco BRL, MD). A 1:2000 dilution of horseradish peroxidase-conjugated secondary antibody (goat anti-rabbit-HRP conjugate, Amersham, IL) was added and incubated for 30 min. The blot was developed with ECL developing reagents (Amersham, IL) and exposed to X-ray film for 10–120 s.

Culture of human umbilical vein endothelial cells (HUVECs)
Segments of fresh human umbilical cords were transported to the laboratory in sterile phosphate-buffered saline. Umbilical veins were cannulated, flushed clear of residual blood, and filled with 0.028% collagenase (Boehringer Mannheim Corp., Indianapolis, IN) in Hanks balanced salt solution (Sigma, St Louis, MO) for 30 min at room temperature. Collagenase solutions containing cells were collected and centrifuged at 220×*g* for 10 min. The cell pellets were resuspended in M199 cell culture medium (Gibco Laboratories, Grand Island, NY) that contained 20% heat-inactivated bovine calf serum (Hyclone Laboratories, Logan, UT) and 20 μg/ml endothelial cell growth factor (Biomedical Technologies Inc., Stoughton, MA). Cell suspensions were incubated in 24-well 50 gelatin-coated tissue culture plates (Becton Dickinson Labware, Lincoln Park, NJ), at 37°C in 5% CO_2 atmosphere. Endothelial cells with characteristic cobblestone morphology usually became confluent in 10–14 days. Experiments were performed with cells that had been passaged only one to two times.

The cells were washed twice with serum-free medium containing 0.1% bovine serum albumin. To duplicate wells of a 24-well tissue culture plate containing confluent HUVECs, 3–15 μg of either boiled or intact streptococcal protease were added, the cells were incubated at 37°C, and monitored periodically by microscopic examination. After 3, 5, and 8 h, 500 μl of 2X SDS–PAGE sample buffer was added to each well. The entire contents of the well (cells plus supernatant medium) were resuspended by repeated pipetting and stored in 1.5 ml Eppendorf tubes at −20°C. The samples were boiled for 5 min and assayed by western immunoblot for ECM proteins as described above. The experiments were performed a total of three times with HUVECs isolated from two different umbilical cords.

Sequencing of the cysteine protease structural gene. The cysteine protease structural gene was amplified by the polymerase chain reaction (PCR), with synthetic oligonucleotides.[4] The oligonucleotide primers used to amplify *speB* and flanking regions were as follows:
SPEB-X, 5′-GTTGTCAGTGTCAACTAACCGT-3′ and
SPEB-2, 5′-ATCTGTGTCTGATGGATAGCTT-3′. The following four oligonucleotides were used as internal sequencing primers:
SPEB-1, 5′-CTTTCTGGCTCTAATATGTATGT-3′;
SPEB-3, 5′-GTTATTGAAAAAGTAAAACC-3′;
SPEB-4, 5′-TTTTCAATAACAGGTGTCAA-3′; and
SPEB-Y, 5′-TCTCCTGAAACGATAACAAA-3′. PCR amplification of 1 μL of chromosomal DNA was performed in 100 μL of a mixture containing 50 mM KCl, 10 mM Tris-HCl, pH 8.3, 1.5 mM $MgCl_2$, 0.001% gelatin, 200 μM each of dATP, dCTP, dGTP, and dTTP, 200 nM each of SPEB-X and SPEB-2, and 2.5 units of *AmpliTaq* DNA polymerase. The thermocycling parameters were denaturation at 94°C for 1 min, annealing at 55°C for 2 min, and extension at 72°C for 2.5 min for a total of 30 cycles. A final extension at 72°C for 15 min was used.

The DNA fragment studied (1437 bp) represents the entire coding region (1197 bp), and 160 bp of upstream and 80 bp of downstream sequence. For about one-third of the strains, single-stranded DNA was prepared by the lambda exonuclease method[64] and sequenced in both orientations with Sequenase version 2.0 (United States Biochemical, OH). Variant alleles were sequenced again to confirm the nucleotide changes.

The protease gene in approximately two-thirds of the strains was characterized by automated DNA sequencing with an Applied Biosystems, Inc., Model 373A instrument. For the automated approach, the gene was amplified with PCR (10 mM Tris-HCl, pH 8.3; 50 mM KCl; 1.5 mM $MgCl_2$; 2.5 units of *Taq* polymerase; 20 picomoles of each primer; 1 μL of chromosomal DNA template), with the following thermocycler parameters: denaturation at 94°C for 4 min, 30 cycles of denaturation at 94°C for 1 min, primer annealing at 55°C for 2 min, extension at 72°C for 2 min, and a final extension at 72°C for 5 min. The unincorporated nucleotides and primers were removed by filtration through Microcon 100 microconcentrators (Amicon Inc., MA). Sequencing reactions with the *Taq* DyeDeoxy terminator cycle sequencing kit (Applied Biosystems, Inc., CA) were performed with 7 μL of PCR amplified DNA as template and 3.2 picomoles of primer. The unincorporated dye terminators and primers were separated from the extension products by spin column purification (Centri-Sep, Princeton Separations, Inc., NJ). The sample was dried in a vacuum centrifuge. Prior to gel loading, the sample was resuspended in 4 μL of sample loading buffer (5:1 deionized formamide:50 mM EDTA, pH 8.0) and heat denatured for 2 min at 90°C. The data were assembled and edited with EDITSEQ, ALIGN, and SEQMAN programs (DNASTAR, WI).

Estimates of genetic relationships among clones. Methods of estimating genetic relationships

among *S. pyogenes* clones by multilocus enzyme electrophoresis have been described.[3] The ET designations are cognate with th se used elsewhere.[3] Thirty-six ETs not identified in a previous study[3] were arbitrarily numbered ET 34–ET 69.

Codon usage in speA, speB, *and* speC. C d n bias for *speA, speB, speC,* and other *S. pyogenes* genes btained from GenBank was calculated with the computer program BIAS provided by T. S. Whittam, Institute for Molecular Evolutionary Genetics, Pennsylvania State University.

Detection of protease expression by immunoblot analysis. Purified protease (100 μg) made by MGAS 1719 mixed with Freund's complete adjuvant was injected subcutaneously at multiple sites to each of two rabbits (Bethyl Labs, TX). Subsequent immunizations with the protease mixed in Freund's incomplete adjuvant were carried out biweekly for a total of five times. Serum was collected and immunoglobulin purified by FPLC with a protein G-Sepharose column (Pharmacia, WI). Specific reactivity with the protease was present in the post- but not pre-immunization serum as assessed by western immunoblot analysis. Bacterial strains were inoculated to 10 ml of a chemically defined medium (JRH Biosciences, PA) and incubated for 10 h at 37°C in 5% CO_2. The cell free supernatants were precipitated with 40% acetone and resuspended in 50 μl PBS (pH 7.4). A total of 3 μg of protein was spotted onto a nylon membrane (Hybond ECL; Amersham, IL), and the unreacted sites blocked by incubation with 0.5% blocking agent (Amersham, IL) for 1 h. The membrane was then rinsed with PBS (pH 7.4)-Tween 20 (0.05%), and incubated for 30 min with purified polyclonal rabbit serum (1:500) directed against the protease. After rinsing the membrane, the secondary antibody (goat anti-rabbit-HRP conjugate, diluted 1:2000) was added, and incubated for an additional 30 min. The blots were developed with ECL developing reagents (Amersham, IL) and exposed to X-ray film.

We thank E. Houston for technical assistance. We are especially grateful for strains recovered from acute rheumatic fever patients supplied by Dr D. LeBlanc from a collection compiled by R. Cole, and Dr D. Bessen from the Lancefield collection at Rockefeller University. We thank Drs S. K. Highlander, R. M. Lebovitz, and J. Sepulveda for suggestions about data presentation, and R. A. Finkelstein for discussions about the *Vibrio cholerae* hemagglutinin/protease. This research was supported by American Heart Association Grant-in-Aid 9200640, and Moran Foundation Project 92–62 (to J. M. M.), and National Institutes of Health Grant HL-47655 (to M. W. J.). J. M. M. is an Established Investigator of the American Heart Association.

References

1. Bisno AL. Group A streptococcal infections and acute rheumatic fever. N Engl J Med 1991; 325: 783–93.
2. Kehoe MA. New aspects of *Streptococcus pyogenes* pathogenicity. Rev Med Microbiol 1991; 2: 147–52.
3. Musser JM, Hauser AR, Kim MH, Schlievert PM, Nelson K, Selander RK. *Streptococcus pyogenes* causing toxic-shock-like syndrome and other invasive diseases: clonal diversity and pyrogenic exotoxin expression. Proc Natl Acad Sci USA 1991; 88: 2668–72.
4. Musser JM, Kapur V, Kanjilal S, Shah U, Musher DM, Barg NL, Johnston KH, Schlievert PM, Henrichsen J, Gerlach D, Rakita RM, Tanna A, Cookson BD, Huang JC. Geographic and temporal distribution and molecular characterization of two highly pathogenic clones of *Streptococcus pyogenes* expressing allelic variants of pyrogenic exotoxin A (scarlet fever toxin). J Infect Dis 1993; 167: 337–46.
5. Elliott SD. A proteolytic enzyme produced by group A streptococci with special reference to its effect on the type-specific M antigen. J Exp Med 1945; 81: 573–92.
6. Elliott SD. The crystallization and serological differentiation of a streptococcal proteinase and its precursor. J Exp Med 1950; 92: 201–19.
7. Liu T-Y, Neumann NP, Elliott SD, Moore S, Stein WH. Chemical properties of streptococcal proteinase and its zymogen. J Biol Chem 1963; 238: 251–6.
8. Gerlach D, Knöll H, Köhler W, Ozegowski J-H, Hrlbalova V. Isolation and characterization of erythrogenic toxins. V. Communication: identity of erythrogenic toxin type B and streptococcal proteinase precursor. Zbl Bakt Hyg, I Abt Orig A 1983; 255: 221–3.
9. Hauser AR, Schlievert PM. Nucleotide sequence of the streptococcal pyrogenic exotoxin type B gene

and relationship between the toxin and the streptococcal proteinase precursor. J Bacteriol 1990; 172: 4536–42.
10. Kellner A, Robertson T. Myocardial necrosis produced in animals by means of crystalline streptococcal proteinase. J Exp Med 1954; 99: 495–504.
11. Björck L, Åkesson P, Bohus M, Trojnar J, Abrahamson M, Olafsson I, Grubb A. Bacterial growth blocked by a synthetic peptide based on the structure of a human proteinase inhibitor. Nature 1989; 337: 385–6.
12. Holm SE, Norrby A, Bergholm A-M, Norgren M. Aspects of pathogenesis of serious group A streptococcal infections in Sweden, 1988–1989. J Infect Dis 1992; 166: 31–7.
13. Tyler SD, Johnson WM, Huang JC, Ashton FE, Wang G, Low DE, Rozee KR. Streptococcal erythrogenic toxin genes: detection by polymerase chain reaction and association with disease in strains isolated in Canada from 1940 to 1991. J Clin Microbiol 1992; 30: 3127–31.
14. Yu C-E, Ferretti JJ. Frequency of the erythrogenic toxin B and C genes (*speB* and *speC*) among clinical isolates of group A streptococci. Infect Immun 1991; 59: 211–5.
15. Kapur V, Majesky MW, Li L-L, Black RA, Musser JM. Cleavage of interleukin 1β precursor to produce active IL-1β by a conserved extracellular cysteine protease from *Streptococcus pyogenes*. Proc Natl Acad Sci USA. 1993; 90: 7676–80.
16. Poon-King R, Bannan J, Viteri A, Cu G, Zabriskie JB. Identification of an extracellular plasmin binding protein from nephritogenic streptococci. J Exp Med 1993; 178: 759–63.
17. Villarreal H Jr, Fischetti VA, van de Rijn I, Zabriskie JB. The occurrences of a protein in the extracellular products of streptococci isolated from patients with acute glomerulonephritis. J Exp Med 1979; 149: 459–72.
18. Jameson BA, Wolf H. The antigenic index: a novel algorithm for predicting antigenic determinants. CABIOS 1988; 4: 181–6.
19. Tai JY, Kortt AA, Liu T-Y, Elliott SD. Primary structure of streptococcal proteinase. III. Isolation of cyanogen bromide peptides: complete covalent structure of the polypeptide chain. J Biol Chem 1976; 251: 1955–9.
20. Yonaha K, Elliott SD, Liu T-Y. Primary structure of streptococcal proteinase. J Protein Chem 1982; 1: 317–34.
21. Bloomster TG, Watson DW. Recent trends in streptococcal pyrogenic exotoxin production by group-A streptococci. In: Holm SE, Christensen P, eds. Basic concepts of streptococci and streptococcal diseases. Surrey: Reedbooks Ltd, 1982; 139–41.
22. Schlievert PM. Role of superantigens in human disease. J Infect Dis 1993; 167: 997–1002.
23. Todd EW. A study of the inhibition of streptococcal proteinase by sera of normal and immune animals and of patients infected with group A hemolytic streptococci. J Exp Med 1947; 85: 591–606.
24. Ruoslahti E. Fibronectin and its receptors. Annu Rev Biochem 1988; 57: 375–413.
25. Yang KD, Bohnsack JF, Hill HR. Fibronectin in host defense: implications in the diagnosis, prophylaxis and therapy of infectious diseases. Pediatr Infect Dis J 1993; 12: 234–9.
26. Simpson WA, Hasty DL, Mason JM, Beachey EH. Fibronectin-mediated binding of group A streptococci to human polymorphonuclear leukocytes. Infect Immun 1982; 37: 805–10.
27. Proctor RA. Fibronectin: an enhancer of phagocyte function. Rev Infect Dis 1987; 9: S412–S419.
28. Rentzsch G, Wilke J. Messung des pH-wertes *in vitro* und *in vivo* an chronisch entzundeten tonsillen. Z Laryng Rhinol Otol 1970; 49: 391–7.
29. Dinarello CA. Interleukin-1 and interleukin-1 antagonism. Blood 1991; 77: 1627–52.
30. Imhoff J-M, Blondeau X, Planchenault T, Emod I, Keil-Dlouha V. Collagenase activation of latent matrix-degrading proteinases from human plasma fibronectin. Biol Chem Hoppe-Seyler 1990; 371: 137–44.
31. Gonzalez-Gronow M, Enghild JJ, Pizzo SV. Streptokinase and human fibronectin share a common epitope: implications for regulation of fibrinolysis and rheumatoid arthritis. Biochim Biophys Acta 1993; 1180: 283–8.
32. Grinnel F, Ho C-H, Wysocki A. Degradation of fibronectin and vitronectin in chronic wound fluid: analysis by cell blotting, immunoblotting, and cell adhesion assays. J Invest Dermatol 1992; 98: 410–6.
33. Finkelstein RA, Boesman-Finkelstein M, Holt P. *Vibrio cholerae* hemagglutinin/lectin/protease hydrolyzes fibronectin and ovomucin: F. M. Burnet revisited. Proc Natl Acad Sci USA 1983; 80: 1092–5.
34. Hase CC, Finkelstein RA. Cloning and nucleotide sequence of the *Vibrio cholerae* hemagglutinin/protease (HA/protease) gene and construction of an HA/protease–negative strain. J Bacteriol 1991; 173: 3311–7.
35. Finkelstein RA, Boesman-Finkelstein M, Chang Y, Hase, CC. *Vibrio cholerae* hemagglutinin/protease, colonial variation, virulence, and detachment. Infect Immun 1992; 60: 472–8.
36. Myhre EB, Kuunsela P. Binding of human fibronectin to group A, C, and G streptococci. Infect Immun 1983; 40: 29–34.
37. Switalski LM, Ljungh A, Ryden C, Rubin K, Hook M, Wadstrom T. Binding of fibronectin to the surface of group A, C and G streptococci isolated from human infections. Eur J Clin Microbiol 1982; 1: 381–7.
38. Talay SR, Valentin-Weigand P, Jerlstrom PG, Timmis KN, Chhatwal GS. Fibronectin-binding protein of *Streptococcus pyogenes*: sequence of the binding domain involved in adherence of streptococci to epithelial cells. Infect Immun 1992; 60: 3837–44.

39. Hanski E, Horwitz PA, Caparon MG. Expression of protein F, the fibronectin-binding protein of *Streptococcus pyogenes* JRS4, in heterologous streptococcal and enterococcal strains promotes their adherence to respiratory epithelial cells. Infect Immun 1992; 60: 5119–25.
40. Hanski E, Caparon MG. Protein F, a fibronectin-binding protein, is an adhesion of the group A streptococcus *Streptococcus pyogenes*. Proc Natl Acad Sci USA 1992; 89: 6172–76.
41. Preissner KT. Structure and biological role of vitronectin. Annu Rev Cell Biol 1991; 7: 275–310.
42. Morihara K, Homma JY. Pseudomonas proteases. In: Holder IA, ed. Bacterial enzymes and virulence. FL: CRC Press, 1985; 41–75.
43. Lantz MS, Allen RD, Vail TA, Switalski LM, Hook M. Specific cell components of *Bacteriodes gingivalis* mediate binding and degradation of human fibrinogen. J Bacteriol 1991; 173: 495–504.
44. Otogoto J-I, Kuramitsu HK. Isolation and characterization of the *Porphyromonas gingivalis prtT* gene, coding for protease activity. Infect Immun 1993; 61: 117–23.
45. Gazzinelli RT, Leme VMC, Cancado JR, Gazzinelli G, Scharfstein J. Identification and partial characterization of *Trypanosoma cruzi* antigens recognized by T cells and immune sera from patients with Chagas' disease. Infect Immun 1990; 58: 1437–44.
46. Murta ACM, Persechini PM, de Souto Padron T, de Souza W, Guimaraes JA, Scharfstein J. Structural and functional identification of GP57/51 antigen of *Trypanosoma cruzi* as a cysteine proteinase. Molec Biochem Parasitol 1990; 43: 27–38.
47. Eakin AE, Mills AA, Harth G, McKerrow JH, Craik CS. The sequence, organization, and expression of the major cysteine protease (cruzain) from *Trypanosoma cruzi*. J Biol Chem 1992; 267: 7411–20.
48. Keene WE, Petitt MG, Allen S, McKerrow JH. The major neutral proteinase of *Entamoeba histolytica*. J Exp Med 1986; 163: 536–49.
49. Keene WE, Hidalgo ME, Orozco E, McKerrow JH. *Entamoeba histolytica*: correlation of the cytopathic effect of virulent trophozoites with secretion of a cysteine proteinase. Exp Parasitol 1990; 71: 199–206.
50. Becker I, Perez-Tamayo R, Montfort I, Alvizouri AM, and Perez-Montfort R. *Entamoeba histolytica*: role of amebic proteinases and polymorphonuclear leukocytes in acute experimental amebiasis in the rat. Exp Parasitol 1988; 67: 268–80.
51. Reed SL, Keene WE, McKerrow JH. Thiol proteinase expression and pathogenicity of *Entamoeba histolytica*. J Clin Microbiol 1989; 27: 2772–7.
52. McKerrow JH, Sun E, Rosenthal PJ, Bouvier J. The proteases and pathogenicity of parasitic protozoa. Annu Rev Microbiol 1993, in press.
53. Bessen D, Jones KF, Fischetti VA. Evidence for two distinct classes of streptococcal M protein and their relationship to rheumatic fever. J Exp Med 1989; 169: 269–83.
54. O'Toole P, Stenberg L, Rissler M, Lindahl G. Two major classes in the M protein family in group A streptococci. Proc Natl Acad Sci USA 1992; 89: 8661–65.
55. Haanes EJ, Heath DG, Cleary PP. Architecture of the *vir* regulons of group A streptococci parallels opacity factor phenotype and M protein class. J Bacteriol 1992; 174: 4967–76.
56. Quentin R, Martin C, Musser JM, Pasquier-Picard N, Goudeau A. Genetic characterization of a cryptic genospecies of *Haemophilus* causing urogenital and neonatal infections. J Clin Microbiol 1993; 31: 1111–6.
57. Selander RK, Musser JM. Population genetics of bacterial pathogenesis. In: Iglewski BH, Clark VL, eds. Molecular basis of bacterial pathogenesis. Orlando FL: Academic Press, 1990; 11–36.
58. Colon-Whitt A, Whitt RS, Cole RM. Production of an erythrogenic toxin (streptococcal pyrogenic exotoxin) by a non-lysogenised group-A streptococcus. In: Parker MT, ed. Pathogenic streptococci. Surrey: Reedbooks Ltd, 1979; 64–5.
59. Johnson LP, Schlievert PM, Watson DW. Transfer of group A streptococcal pyrogenic exotoxin production to nontoxigenic strains by lysogenic conversion. Infect Immun 1980; 28: 254–7.
60. Nida SK, Ferretti JJ. Phage influence on the synthesis of extracellular toxins in group A streptococci. Infect Immun 1982; 36: 745–50.
61. Cowe E, Sharp PM. Molecular evolution of bacteriophages: discrete patterns of codon usage in T4 genes are related to the time of gene expression. J Mol Evol 1991; 33: 13–22.
62. Post LE, Strycharz GD, Nomura M, Lewis H, Dennis PP. Nucleotide sequence of the ribosomal protein gene cluster adjacent to the gene for RNA polymerase subunit β in *Escherichia coli*. Proc Natl Acad Sci USA 1979; 76: 1697–1701.
63. Bennetzen JL, Hall BD. Codon selection in yeast. J Biol Chem 1982; 257: 3026–31.
64. Grantham R, Gautier C, Gouy M, Mercier R, Pavé A. Codon catalog usage and the genome hypothesis. Nucl Acids Res 1980; 8: r49–r62.
65. Sharp PM, Li W-H. The codon adaption index—a measure of directional synonymous codon usage bias, and its potential applications. Nucl Acids Res 1987; 15: 1281–95.
66. Sharp PM, Li W-H. The rate of synonymous substitution in enterobacterial genes is inversely related to codon usage bias. Mol Biol Evol 1987; 4: 222–30.
67. Fischetti VA. Streptococcal M protein: molecular design and biological behavior. Clin Microbiol Rev 1989; 2: 285–314.
68. Chen CC, Cleary PP. Complete nucleotide sequence of the streptococcal C5a peptidase gene of *Streptococcus pyogenes*. J Biol Chem 1990; 265: 3161–7.
69. O'Connor SP, Darip D, Fraley K, Nelson CM, Kaplan EL, Cleary PP. The human antibody response to streptococcal C5a peptidase. J Infect Dis 1991; 163: 109–16.

70. Cleary PP, Handley J, Suvorov AN, Podbielski A, Ferrieri P. Similarity between the group B and A streptococcal C5a peptidase genes. Infect Immun 1992; 60: 4239–44.
71. Rhim SH, Millar SE, Robey F, Luo A-M, Lou Y-H, Yule T, Allen P, Dean J, Tung KSK. Autoimmune disease of the ovary induced by a ZP3 peptide from the mouse zona pellucida. J Clin Invest 1992; 89: 28–35.
72. McCarty M. The occurrence of nucleases in culture filtrates of group A hemolytic streptococci. J Exp Med 1948; 88: 181–8.
73. Higuchi RG, Ochman H. Production of single-stranded DNA templates by exonuclease digestion following the polymerase chain reaction. Nucl Acids Res 1989; 17: 5865.